1949

Sequential tests of the linear hypothesis

Osmer S. Carpenter
Iowa State College

SEQUENTIAL TESTS OF THE LINEAR HYPOTHESIS

by

Osmer S. Carpenter

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State College

1949

TABLE OF CONTENTS

I. INTRODUCTION

A. Definition and Explanation of the Problem

The general normal linear hypothesis can be defined as follows:

Let x_i: $i = 1, \ldots n$, be normally and independently distributed random variates with mean ξ_i and common variance σ^2, where ξ_i: $i = 1, \ldots, n$ are constrained to the p-dimensional linear subspace defined by $n - p$ independent linear restrictions on the ξ_i. It is required to test the hypothesis that ξ_i, $i = 1, \ldots, n$ is a point of the $p - r$ dimensional sub-space defined by $\underline{r}$ additional independent linear restrictions. It is often convenient to define the $n - p$ given linear restrictions by expressing each ξ_i as a linear function of $\underline{p}$ unknown constant parameters and to state the r additional restrictions of the hypothesis by specifying fixed values for $\underline{r}$ of these parameters. In this form it is possible to consider an alternative assumption in which some or all of the parameters are assumed to be independent random variates and the hypothesis to be tested is that some $\underline{r}$ of these parameters have zero variance.

The linear hypothesis, with either constant or variable parameters, occurs frequently in statistics and forms the mathematical basis for most analysis of variance and regression problems as well as for the theory of design of experiments. The theory and application of tests of the linear hypothesis have been the object of considerable investigation and extensive literature on the subject exists. With a few exceptions, available tests are based on fixed sample size. The newly developed technique of sequential analysis makes available, for some statistical tests, a method in which sample size is determined by the outcome of the observations, resulting on the average in fewer observations for a required test strength. The object of the present study is to develop methods for the application of sequential analysis to the general normal linear hypothesis.

B. History of the Problem

Large sample methods for testing certain normal linear hypotheses were in use prior to 1900. In 1908, the development of the t-test by Student provided the first application of small sample theory to a problem of this type and in the next twenty-five years the development of the analysis of variance and the use of the variance ratio test led to a fairly complete solution of the general problem. A large part of this advance was due to the work of R. A. Fisher. Since 1930 a large number of contributions have been made to the theory and applications of these methods. The methods of application are discussed in a number of texts, in particular those of Fisher (13), Snedecor (34) and Goulden (17); more complex examples are to be found in a large variety of sources. A large amount of work has also been done on the theoretical aspects of the problem but no comprehensive treatment of the subject has been published.

One of the first treatments of the mathematical aspects of the problem was by J. O. Irwin in 1931 (23) who discussed the mathematical assumptions involved. St. Kolodziеczyk in 1935 (27) applied the technique of the likelihood ratio test to the problem, and showed that the likelihood ratio test is equivalent to the usual variance ratio test. In 1938, P. C. Tang (39) discussed the power function of the usual tests, proved that the power function depended on a single parameter, λ, and prepared tables of this function. Later, Hsu (21) and Wald (40) showed that the test is uniformly most powerful of all tests having power functions dependent on the parameter, λ, only. Recently, the problem of most powerful tests of the linear hypothesis has been discussed by E. H. Lehmann and C. Stein (28).

The theory and application of sequential tests were developed largely by the work of A. Wald, first published in non-restricted form in 1945 (42) and more completely in book form in 1947 (46). Most of present literature on this subject deals with the theory and application of Wald's sequential probability ratio test. For a test of a simple hypothesis against a single-valued alternative and in particular where successive observations are independent, a rather complete theory exists. Methods have also been devised for dealing with composite hypotheses and multi-valued alternatives. Previous applications of sequential analysis to testing of linear hypothesis include a sequential t-test and a test that the means of a number of classes are equal to specified constant values when the variance is known, both of which are discussed by Wald (46). A test of whether the variance of one population is less or greater than the variance of a second population has been discussed by Girshick (15).

C. Plan of Investigation

The present discussion consists of three parts. The first portion consists of the review, consolidation and extension of a number of previous results concerning the form of the linear hypothesis, the construction of test criteria and their distribution under various alternatives. The second section gives a brief review of the theory and application of sequential tests. The third part discusses the construction and characteristics of several possible sequential tests of the general normal linear hypothesis.

II. REVIEW AND EXTENSION OF PREVIOUS RESULTS

The literature on the theory of the general linear hypothesis is to be found in a number of widely scattered sources. The purpose of this section is to bring together some previously known results which are essential to the present problem.

A. The General Linear Hypothesis

Before constructing a sequential test of the linear hypothesis, it is necessary to review and consolidate some of the known results on the form of the usual test functions and their distributions. The distribution of the ratio of quadratic forms under various hypotheses receives particular consideration as it is basic to all common tests and is especially important in the sequential case.

1. Specification of the general linear hypothesis.

The general linear hypothesis in normal variation can be defined as follows:

Let $x_1, \ldots, x_n$ be a random sample of independent variates where x_i has the normal distribution

$$(1) \qquad f(x_i) = (1/\sqrt{2\pi}\,\sigma)\, e^{-(x_i - \xi_i)^2/2\sigma^2}$$

Let the general parameter space Ω consist of the $n + 1$ dimensional space of the ξ_i and σ, $\Omega : [\xi_1, \ldots, \xi_n, \sigma]$, restricted by $n - p$ linearly independent restrictions on the ξ_i. The hypothesis to be tested is that the true parameter point is an element of ω, where $\omega \in \Omega$ is specified by $r \leq p$ additional independent linear restrictions.

The linear restrictions of Ω and ω can be specified in various forms, including the following, with appropriate conditions for linear independence in each case:

a. Homogeneous linear restrictions on the ξ_i.

$$(2) \qquad \Omega : \sum_{j=1}^{n} a_{ij}\,\xi_j = 0, \qquad i = 1, \ldots, n - p$$

$$\omega : \sum_{j=1}^{n} b_{ij}\,\xi_j = 0, \qquad i = n - p + 1, \ldots, n - p + r$$

b. Restrictions expressed in terms of parameters.

$$(3) \qquad \Omega : \xi_i = \sum_{j=1}^{p} c_{ij}\,\theta_j \qquad i = 1, \ldots, n$$

$$\omega : \qquad \sum_{j=1}^{p} e_{ij}\,\theta_j = 0, \quad i = 1, \ldots, r$$

c. <u>Canonical form with parameters</u>.

$$(4)\quad \Omega : \xi_i = \sum_{j=1}^{p} c_{ij}\,\theta_j \qquad i = 1, \ldots, n$$

$$\omega : \quad \theta_j = 0 \qquad j = p - r + 1, \ldots, p$$

Though in practice it usually will be unnecessary, it is possible to change from one form of hypothesis to another by transformation of parameters. It can be seen that form (c) is a particular case of (b). The general form (b) can be reduced to (c) by using the parameter transformation $\theta_i' = \sum_{j=1}^{p} e_{ij}\theta_j$, $(i = 1, \ldots, r)$, solving for <u>r</u> of the θ_j and substituting in the equations of Ω. The existence of a transformation from (a) to (b) can be shown as follows: Let $A = (a_{ij})$ and $B = (b_{ij})$ be the matrices of coefficients of the linear restrictions, arranged so that $\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{pmatrix}$, where A_1 and B_2 are of orders n-p and r and non-singular. Let $\xi_1 = (\xi_1, \ldots, \xi_{n-p})$, $\xi_2 = (\xi_{n-p+1}, \ldots, \xi_{n-p+r})$, $\xi_3 = (\xi_{n-p+r+1}, \ldots, \xi_n)$ be column vectors. Let $A\xi = 0$. Then $\xi_1 = -A_1^{-1} A_2 \xi_2 - A_1^{-1} A_3 \xi_3$. Let $\begin{pmatrix} \xi_2 \\ \xi_3 \end{pmatrix} = K\theta$, where K is of order p and non-singular. Then $\xi_1 = A_1^{-1} (A_2 \vdots A_3)K\theta$ and thus $\xi_i = \sum_{j=1}^{p} c_{ij}\theta_j$, where $C = \begin{pmatrix} A_1^{-1} (A_2 \vdots A_3) \\ \hdashline I \end{pmatrix} K$ and is of rank p. In ω:

$B\xi = BC\theta$, where

$$BC = (B_1 \mid B_2 \mid B_3) \begin{pmatrix} -A_1^{-1} A_2 & -A_1^{-1} A_3 \\ I & 0 \\ 0 & I \end{pmatrix} K$$

$$= (-B_1A_1^{-1}A_2 + B_2 \mid -B_1A_1^{-1}A_3 + B_3)\ K.$$

$$\text{But } (-B_1A_1^{-1} \mid I) \begin{pmatrix} A_1 & A_2 \\ B_1 & B_2 \end{pmatrix}$$

$$= (0 \mid -B_1A_1^{-1} A_2 + B_2)$$

which is of rank _r_, since the first factor has rank _r_ and the second is non-singular; thus BC is of rank _r_.

The form (c) will be used for most of this discussion. It will sometimes be expressed in the equivalent form

$x_i = \sum_{j=1}^{p} c_{ij}\theta_j + e_i$, where e_i is $N(0, \sigma^2)$. For the purpose of mathematical theory it will sometimes be desirable to make use of the following theorem adapted from a more general theorem due to P. L. Hsu (20).

Theorem I:

Let x_i, $i = 1, \ldots, n$, be normally and independently distributed with means ξ_i and common unknown variance σ^2 under the linear restrictions:

$$(5) \quad \Omega: \quad \xi_i = \sum_{j=1}^{p} c_{ij}\theta_j, \qquad i = 1, \ldots, n,$$

$$\omega: \quad \theta_j = 0, \qquad j = 1, \ldots, r.$$

Then there exists a real orthogonal linear transformation $z = \Gamma x$, such that the z_i are normally and independently distributed, with variance σ^2 and $E(z_i) = \eta_i$, under the restrictions

$$(6) \quad \begin{aligned} \Omega &: \eta_i = 0, \quad i = p+1, \ldots, n \\ \omega &: \eta_i = 0, \quad i = p-r+1, \ldots, p. \end{aligned}$$

Proof: Let $X = (x_1, \ldots, x_n)$ and $\theta = (\theta_1, \ldots, \theta_n)$ be column vectors and let C be the n x p matrix of (c_{ij}). Since C is of rank p, $C'C$ is a positive definite matrix of order p and therefore can be expressed as the product EE' where E is a triangular matrix

$$(7) \quad E = \begin{pmatrix} e_{11} & e_{12} & \cdots\cdots & e_{1n} \\ 0 & e_{22} & \cdots\cdots & e_{2p} \\ \cdot & \cdot & \cdots\cdots & \cdot \\ 0 & \cdot & \cdots\cdots 0 & e_{pp} \end{pmatrix}, \quad e_{ii} \neq 0.$$

Let $\Gamma_1 = E^{-1}C'$, a $\underline{p}$ times $\underline{n}$ matrix of rank $\underline{p}$.
Then $\Gamma_1 \Gamma_1' = E^{-1}C'CE^{-1\prime} = I$. Let Γ_1 be the first p rows of a real orthogonal matrix Γ of order n, $\Gamma = \begin{pmatrix} \Gamma_1 \\ \Gamma_2 \end{pmatrix}$, and let $Z_1 = (z_1, \ldots, z_p)$, $Z_2 = (z_{p+1}, \ldots, z_n)$ be column vectors. Make the transformation $Z = \Gamma X$. Since the transformation is orthogonal, the variates z_i are independent with common variance σ^2. Let $E(z_i) = \eta_i$. Then

$$(8) \qquad \eta = E(z \mid c) = \Gamma ce = \begin{pmatrix} \Gamma_1 ce \\ \cdots \\ \Gamma_2 ce \end{pmatrix},$$

$$\Gamma_1 ce = \Gamma_1 \Gamma_1' E' e = E' e,$$

$$\Gamma_2 ce = \Gamma_2 \Gamma_1' E' e = 0.$$

Therefore $(z_1, \ldots, z_n)$ are normally and independently distributed, with variance σ^2, means $\eta_i = \sum_{j=1}^{i} e_{ji} e_j, (i = 1, \ldots, p)$ and $\eta_i = 0, (i = p + 1, \ldots, n)$. The restrictions $\omega: e_i = 0,$ $(i = 1, \ldots, r)$ are equivalent to $\eta_i = 0, (i = 1, \ldots, r)$ and by interchange of subscripts we have the given theorem. It is apparent that the other forms for the linear hypothesis are reducible to this form.

2. Distribution functions related to the linear hypothesis.

Before considering the actual test of the normal linear hypothesis, it is useful to consider the distribution of various statistics related to the normal distribution.

a. χ^2 distribution.

Let

$$(9) \qquad p(\chi^2, r, \lambda^2) = \tfrac{1}{2} e^{-\lambda^2} (\chi^2/2)^{\frac{r}{2}-1} e^{-\chi^2/2} \sum_{i=0}^{\infty} \frac{(\lambda^2 \chi^2/2)^i}{i!\,\Gamma(\frac{r}{2}+i)}$$

$r > 0, \ \lambda^2 \geqslant 0, \ 0 \leqslant \chi^2 < \infty$.

The characteristic function of this distribution is given by

$$(10) \qquad \varphi(t) = E e^{it\chi^2} = e^{2it\lambda^2/(1-2it)} (1-2it)^{-r/2}$$

and the lower moments can be evaluated from the cumulative function,

$$\text{(11)} \qquad \ln \varphi(t) = \sum_{n=1}^{\infty} n!\, 2^n(\lambda^2 + r/2n)\,(it)^n/n!$$

$$\gamma_n = 2^{n-1}\,(n-1)!\,(2n\lambda^2 + r)$$

$$\gamma_1 = 2\lambda^2 + r, \quad \gamma_2 = 2(4\lambda^2 + r),$$

$$\gamma_3 = 8(6\lambda^2 + r), \quad \gamma_4 = 48(8\lambda^2 + r).$$

The χ^2 distribution is additive, for if χ_1^2 and χ_2^2 are independently distributed by $p(\chi_1^2, r_1, \lambda_1^2)$ and $p(\chi_2^2, r_2, \lambda_2^2)$, then

$$\text{(12)} \qquad \varphi(t) = E\, e^{it(\chi_1^2 + \chi_2^2)} = e^{2it(\lambda_1^2 + \lambda_2^2)/(1-2it)} \cdot (1-2it)^{-(r_1 + r_2)/2}$$

which is the characteristic function of the distribution $p(\chi_1^2 + \chi_2^2;\, r_1 + r_2,\, \lambda_1^2 + \lambda_2^2)$. Now, let x be distributed normally with mean ξ and unit variance and let $s = x^2$. Then

$$\text{(13)} \qquad p(s) = (8\pi s)^{-1/2}\left[e^{-(\sqrt{s}-\xi)^2/2} + e^{-(\sqrt{s}+\xi)^2/2}\right]$$

$$= (1/2)e^{-\lambda^2}(s/2)^{-1/2}\, e^{-s/2} \sum_{i=0}^{\infty} \frac{(\lambda^2 s/2)^i}{i!\,\Gamma(1/2 + i)}$$

and hence x^2 is distributed by $p(\chi^2;\, 1, \lambda^2)$, where $\lambda^2 = \xi^2/2$. If $\underline{x}$ is $N(\xi, \sigma^2)$ then x^2/σ^2 is distributed by $p(\chi^2;\, 1, \lambda^2)$ where $\lambda^2 = \xi^2/2\sigma^2$ and if $x_1, \ldots, x_n$ are distributed normally and independently with means ξ_i and common variance σ^2, it follows from the additive property that $\sum x_i^2/\sigma^2$ is distributed as $p(\chi^2;\, n, \lambda^2)$ where $\lambda^2 = \sum \xi_i^2/2\sigma^2$.

b. <u>Distribution of the ratio of two independent χ^2 variates.</u>

Next consider the distribution of the ratio of two independent χ^2 variates. Let χ_1^2 and χ_2^2 be independently distributed by $p(\chi_1^2, r_1, \lambda_1^2)$ and $p(\chi_2^2, r_2, \lambda_2^2)$ and let $u = \chi_1^2/\chi_2^2$, $v = \chi_2^2$. Then

$$(14) \qquad p(u) = e^{-(\lambda_1^2+\lambda_2^2)} \int_0^\infty \Big\{ u^{r_1/2-1} (v/2)^{(r_1+r_2)/2-1} e^{-(1+u)v/2} \sum_{i=0}^\infty \frac{(\lambda_1^2 u)^i \ (v/2)^i}{i!\,\Gamma(r_1/2+i)} \sum_{j=0}^\infty \frac{(\lambda_2^2\, v/2)^j}{j!\,\Gamma(r_2/2+j)} \Big\} dv/2$$

$$= e^{-(\lambda_1^2+\lambda_2^2)}\, u^{r_1/2-1}\,(1+u)^{-(r_1+r_2)/2} \sum_{i=0}^\infty \sum_{j=0}^\infty \frac{\lambda_2^{2j}\,(\lambda_1^2 u)^i}{i!\,j!\,(1+u)^{i+j}}\, B^{-1}(r_1/2+i,\ r_2/2+j),$$

$0 \leq u < \infty$.

In the particular case $\lambda_2^2 = 0$, $\lambda_1^2 = \lambda^2$,

$$(15) \qquad p(u;\, r_1,\, r_2,\, \lambda^2) = e^{-\lambda^2} u^{r_1/2-1} (1+u)^{-(r_1+r_2)/2} \sum_{i=0}^\infty \frac{(\lambda^2 u)^i (1+u)^{-i}}{i!\,B\,(r_1/2+i,\ r_2/2)}$$

or in terms of $w = u/(1+u)$,

$$(16) \qquad p(w;\, r_1,\, r_2,\, \lambda^2) = e^{-\lambda^2} w^{r_1/2-1} (1-w)^{r_2/2-1} \sum_{i=0}^\infty \frac{(\lambda^2 w)^i}{i!\,B(r_1/2+i,\ r\ /2)}$$

$$= \frac{e^{-\lambda^2} w^{r_1/2 - 1}(1 - w)^{r_2/2 - 1}}{B(r_1/2,\ r_2/2)}$$

$$F\left[(r_1 + r_2)/2,\ r_1/2,\ \lambda^2 w\right]$$

where

(17) $$F(a,\ b,\ tx) = 1 + \frac{a}{b} \cdot \frac{tx}{1!} + \frac{a(a+1)}{b(b+1)} \cdot \frac{(tx)^2}{2!} + \ldots$$

is the confluent hypergeometric function. By Kummer's Formula, $F(a + b,\ a,\ tx) = e^{tx} F(-b,\ a,\ -tx)$, we can write the alternative form

(18) $$p(w;\ r_1,\ r_2,\ \lambda^2) = B^{-1}(r_1/2,\ r_2/2) e^{-\lambda^2} w^{r_1/2 - 1}$$

$$(1 - w)^{r_2/2 - 1} e^{\lambda^2 w} F\left[-r_2/2,\ r_1/2, -\lambda^2 w\right]$$

which is a finite series when r_2 is even. Properties of the $\underline{u}$ distribution have been discussed by J. Wishart (51). The following properties of $F(a,\ B,\ tx)$ and $p(w)$ are noted for reference:

(1) $F(a,\ b,\ tx)$ is a monotonic increasing function of x; $a,\ b,\ t,\ x > 0$. For

(19) $$\partial F/\partial x = \frac{a}{b} \cdot \frac{t}{1!} + \frac{a(a+1)}{b(b+1)} \cdot \frac{t^2 x}{1!} + \frac{a(a+1)(a+2)}{b(b+1)(b+2)} \cdot \frac{t^3 x^2}{2!} + \cdots$$

$$= (at/b)\ F(a + 1,\ b + 1,\ tx) > 0\ .$$

(20) $$\partial^2 F/\partial x \partial t = (a/b)\left[F(a + 1,\ b + 1,\ tx)\right.$$

$$\left. + \frac{t^2(a+1)}{b+1} F(a + 2,\ b + 2,\ tx)\right] > 0.$$

(2) $\text{Prob}(w > w_0 \mid \lambda^2)$ is an increasing function of λ^2. For

(21) $$\text{Prob}(w > w_0 \mid \lambda^2) = 1 - \int_0^{w_0} p(w; r_1, r_2, \lambda^2)\, dw,$$

$$\text{Prob}(w < w_0 \mid \lambda^2) = \int_0^{w_0} p(w; r_1, r_2\, \lambda^2)\, dw$$

$$= e^{-\lambda^2} \sum_{i=0}^{\infty} (\lambda^{2i}/i!)\, I(r_1/2 + i,\, r_2/2,\, w_0)$$

where $I(a, b, x)$ is the incomplete B-function,

(22) $$B^{-1}(a, b,) \int_0^x t^{a-1}(1 - t)^{b-1} dt.$$

(23) $$d\text{Prob}(w < w_0 \mid \lambda^2)/d\lambda^2 = e^{-\lambda^2} \sum_{i=0}^{\infty} (\lambda^{2i}/i!)$$

$$\left\{ I\left[(r_1/2) + i + 1,\, r_2/2,\, w_0\right] - I\left[(r_1/2) + i,\, r_2/2, w_0\right] \right\}$$

But

(24) $$I(a + 1, b, x) - I(a, b, x) = B^{-1}(a + 1, b) \int_0^x t^a (1 - t)^{b-1} dt$$

$$- B^{-1}(a, b) \int_0^x t^{a-1}(1 - t)^{b-1}\, dt$$

$$= B^{-1}(a, b) \left\{ \left[(a + b)/a\right] \int_0^x t^a (1 - t)^{b-1}\, dt \right.$$

$$\left. - \int_0^x t^{a-1}\, (1 - t)^{b-1}\, dt \right\},$$

(25) $$\int_0^x t^a (1 - t)^{b-1}\, dt = - x^a (1 - x)^b/(a + b)$$

$$+ \left[a/(a + b)\right] \int_0^x t^{a-1}\, (1 - t)^{b-1}\, dt.$$

Hence

(26) $$I(a + 1, b, x) - I(a, b, x) = - x^a (1 - x)^b / aB(a, b) < 0\,.$$

c. <u>The distribution of quadratic forms.</u>

We next consider some theorems connected with the distribution of quadratic forms in normal variates, the first being an extension of a theorem due to A. T. Craig (9) and H. Hotelling (19).

Theorem II: Let $x_1, \ldots, x_n$ be a set of normally and independently distributed variates with equal variance σ^2 and means $m_1, \ldots, m_n$. Let $Q_1(x) = X'A_1X$ and $Q_2(x) = X'A_2X$ be symmetric quadratic forms of ranks r_1 and r_2 in X. Then a necessary and sufficient condition that Q_1 and Q_2 be statistically independent for all values of M is that $A_1A_2 = 0$.

Further, a necessary and sufficient condition that a symmetric $Q(x) = X'AX$ of rank $\underline{r}$ be distributed as $\sigma^2\chi^2$ with $\underline{r}$ degrees of freedom is that $A^2 = A$. If $Q(x)/\sigma^2$ is distributed by $p(\chi^2, r, \lambda^2)$ then $\lambda^2 = Q(m)/2\sigma^2$.

Proof: In the following, the A, A_1, A_2 are symmetric matrices, L is an orthogonal matrix, Γ is a diagonal matrix of characteristic roots and X, Y, M, U are column vectors.

Consider the moment generating function of the χ^2 distribution,

(27) $$M(t) = Ee^{t\chi^2/2} = e^{\lambda^2 t/(1-t)}(1-t)^{-r/2} .$$

Let us derive the corresponding function for a quadratic form Q(x) and show that the condition $A^2 = A$ is necessary and sufficient for their equality. Let $x_1, \ldots, x_n$ be normally and independently distributed with $E(x_i) = m_i$ and $\sigma^2 = 1$. By direct transformation or by moment generating functions it can be shown that

(28) $$p(x_i^2) = p(\chi^2, 1, \lambda_i^2), \text{ where } \lambda_i^2 = m_i^2/2 .$$

Using a well-known property of moment generating functions,

(29) $$Ee^{(t/2)\sum_{i=1}^{n}\alpha_i x_i^2} = \prod_{i=1}^{n} Ee^{(t/2)\alpha_i x_i^2}$$

$$= e^{\sum t\alpha_i \lambda_i^2/(1-t\alpha_i)} \prod(1-t\alpha_i)^{-1/2}.$$

Let $Q(x) = X'AX$ be a real symmetric quadratic form of rank r. Then there exists an orthogonal linear transformation L such that the transformation $X = LY$ gives $Q = \sum \gamma_i \; y_i^2$, or in matrix notation $X'AX = Y'L'ALY = Y'\Gamma Y$, where γ_i are the characteristic roots of A. Hence

(30) $$G_Q(t) = Ee^{tQ/2} = \prod(1-t\gamma_i)^{-1/2} \exp(1/2)\sum u_i^2 \; \gamma_i/(1-t\gamma_i)$$

$$= |I - tA|^{-1/2} \exp\left\{-(1/2)M'tA(I-tA)^{-1} M\right\} ,$$

where $u_i = E(y_i)$, $U = L'M$, and $\underline{t}$ is restricted to values for which $|I - tA| \neq 0$. The change in form of the exponential can be shown as follows:

(31) $$\sum u_i^2 \gamma_i \; t/(1-t\gamma_i) = U'\left[I - (I - t\Gamma)^{-1}\right] U$$

$$= M'L\left[I - (I - t\Gamma)^{-1}\right] L'M$$

$$= M'\left[I - (I - tA)^{-1}\right] M$$

$$= -M'tA(I - tA)^{-1} M.$$

A necessary and sufficient condition that $G_Q(t) = M(t)$ is that $A^2 = A$. If $A^2 = A$, then all of the characteristic roots of A are $+1$ or 0 and the sufficiency of the condition can be established by substituting the correct values of $\gamma_i = +1$ or 0 into the first form of $G_Q(t)$ giving

$$(32) \qquad G_Q(t) = e^{t\lambda^2/(1-t)}(1-t)^{-r/2} = M(t)$$

where $\lambda^2 = \sum \gamma_i u_i^2/2$. Thus the condition is sufficient. Since $\sum \gamma_i u_i^2 = U'\Gamma U = M'L\Gamma L'M = M'AM = Q(m)$, it follows that $\lambda^2 = Q(m)/2$. If $G_Q(t) = M(t)$ identically in M, let $M = 0$, reducing the required condition to $|I - tA|^{-1/2} = (1-t)^{-r/2}$. It has been shown by Craig that a necessary (and sufficient) condition for this equality is that $A^2 = A$. Hence this condition is also necessary in the original identity.

Next, let $Q_1 = X'A_1X$ and $Q_2 = X'A_2X$ be real symmetric quadratic forms of ranks r_1 and r_2. Then from equation (30),

$$(33) \qquad G(t_1, t_2) = Ee^{t_1Q_1/2 + t_2Q_2/2}$$

$$= e^{-(1/2)M'(t_1A_1 + t_2A_2)(I-t_1A_1-t_2A_2)^{-1}}$$

$$\cdot|I-t_1A_1-t_2A_2|^{-1/2},$$

t_1 and t_2 being restricted to values for which $(I-t_1A_1-t_2A_2)$, $(I-t_1A_1)$, and $(I-t_2A_2)$ are all non-singular. A necessary and sufficient condition that $G(t_1,t_2) = G_Q(t_1)G_Q(t_2)$ is $A_1A_2 = 0$.

The required condition is

$$G(t_1, t_2) = e^{-(1/2)M' t_1 A_1 (I - t_1 A_1)^{-1} M} \left| I - t_1 A_1 \right|^{-1/2} \cdot e^{-(1/2)M' t_2 A_2 (I - t_2 A_2)^{-1} M} \left| I - t_2 A_2 \right|^{-1/2} \tag{34}$$

Assume $A_1 A_2 = 0$. Then $(I - t_1 A_1 - t_2 A_2) = (I - t_1 A_1)(I - t_2 A_2)$ and $|I - t_1 A_1 - t_2 A_2| = |I_1 - t_1 A_1| \cdot |I - t_2 A_2|$. Also $(t_1 A_1 + t_2 A_2)(I - t_1 A_1 - t_2 A_2)^{-1} = t_1 A_1 (I - t_1 A_1)^{-1} + t_2 A_2 (I - t_2 A_2)^{-1}$; for using the identity $tA(I - tA)^{-1} = (I - tA)^{-1} - I$, this becomes $(I - t_2 A_2)^{-1}(I - t_1 A_1)^{-1} = (I - t_1 A_1)^{-1} + (I - t_2 A_2)^{-1} - I$, and by multiplying both sides on the left by $I - t_2 A_2$ and on the right by $I - t_1 A_1$ the identity follows. Thus the condition is sufficient.

If $G(t_1, t_2) \equiv G_Q(t_1)\, G_Q(t_2)$ identically in M, then let $M = 0$ and the equation reduces to that discussed by Hotelling, who showed that a necessary condition that $|I - t_1 A_1 - t_2 A_2| = |I - t_1 A_1| \cdot |I - t_2 A_2|$ is that $A_1 A_2 = 0$. Therefore this condition is also necessary.

It is also appropriate to note a theorem due to Craig, that if a series of quadratic forms $Q_1, \ldots, Q_k$ are such that $A_1 + \ldots + A_k = I$, then the independence condition $A_i A_j = 0$, $i \neq j$, implies the χ^2 distribution condition $A_i^2 = A_i$.

A second useful theorem on the independence and form of distribution of quadratic forms is the following due to Cochran (7) and extended to the non-central case by Madow (30).

Theorem III: Let $Q_1, \ldots, Q_m$ be a set of quadratic forms in $x_1, \ldots, x_n$ with ranks $r_1, \ldots, r_m$ and such that $\sum_{i=1}^{n} x_i^2 = \sum_{j=1}^{m} Q_j$. Then a necessary and sufficient condition that there exist a single orthogonal linear transformation $X = LY$ such that

$$Q_1 = \sum_{1}^{r_1} y_i^2, \quad Q_2 = \sum_{r_1+1}^{r_1+r_2} y_i^2, \ldots, \tag{35}$$

$$Q_m = \sum_{r_1+\ldots+r_{n-1}+1}^{r_1+\ldots+r_m} y_i^2$$

is that $\sum_{j=1}^{m} r_j = n$.

If the x_i are normally and independently distributed with means m_i and common variance σ^2, it follows from the existence of the above transformation that each Q_j/σ^2 is independently distributed by $p(\chi^2; r_j, \lambda_j^2)$, where $\lambda_j^2 = Q_j(m)/2\sigma^2$.

d. <u>The distribution of the ratio of two independent quadratic forms.</u>

From the above distributions, it is now a simple matter to derive the distribution of the ratio of two independent quadratic forms. Let $Q_1(x)$ and $Q_2(x)$ be independent quadratic forms distributed as $\sigma_1^2\chi^2$ and $\sigma_2^2\chi^2$ with degrees of freedom r_1 and r_2, and parameters λ_1^2 and λ_2^2. Let $v = Q_1/Q_2$, and

$K = \sigma_1^2/\sigma_2^2$. Then V/K is distributed as $p(u; r_1, r_2, \lambda_1^2, \lambda_2^2)$ given by equation (14). In the particular cases $\lambda_2^2 = 0$, $\lambda_1^2 = \lambda^2$ and $K = 1$, $p(v) = P(u;, r_1, r_2, \lambda^2)$ given by equation (15) and $p[(v/(1 + v)] = p(w; r_1, r_2, \lambda^2)$ equation (16). In the case $\lambda_1^2 = \lambda_2^2 = 0$, $K \neq 1$,

$$(36) \qquad P(v) = K^{r_2/2} B^{-1}(r_1/2, r_2/2)\, v^{r_1/2 - 1} (K + v)^{-(r_1 + r_2)/2}$$

where $0 < v < \infty$ when $K > 0$.

3. Tests of the general linear hypothesis.

In testing the general linear hypothesis, it is necessary to consider two different assumptions regarding the parameters θ_j: (1) The parameters θ_j are unknown constants (2) Some or all of the θ_j are random variables. Let us consider the basic test functions and their distributions under these two assumptions.

a. Quadratic forms basic to the desired tests.

Let x_i be distributed normally and independently, $N(\xi_i, \sigma^2)$, where $\xi_i = \sum_{j=1}^{p} c_{ij}\theta_j$, $i = 1, \ldots, n$, and let it be required to test the hypothesis H_0: $\theta_j = 0$, $j = p - r + 1, \ldots, p$. Let

$$(37) \qquad S^2 = \sum_i (x_i - \xi_i)^2 = \sum_i (x_i - \sum_j c_{ij}\theta_j)^2$$

Let S_a^2 be the minimum of S^2 in the general $p + 1$ dimensional parameter space Ω: $-\infty < \theta_j < \infty$, $0 < \sigma^2 < \infty$, that is

(38) $$S_a^2 = \sum_i (x_i - \sum_j c_{ij}\hat{\theta}_j)^2 ,$$

where $\hat{\theta}_j$ is the least square estimate of θ_j in Ω. Similarly, let S_r^2 be the minimum of S^2 in the $p + 1 - r$ dimensional parameter sub-space ω restricted by H_o,

(39) $$S_r^2 = \sum_i (x_i - \sum_j c_{ij}\hat{\hat{\theta}}_j)^2$$

where $\hat{\hat{\theta}}_j$, $j = 1, \ldots, p - r$, $\hat{\hat{\theta}}_j = 0$, $j = p - r + 1, \ldots, p$ are the least square estimates of θ_j in ω. Let

(40) $$S_b^2 = S_r^2 - S_a^2 .$$

b. <u>The distribution of S_a^2, S_r^2, S_b^2 and related ratios.</u>

Theorem IV: In the case where the θ_j are constant parameters, S_a^2/σ^2 is distributed by $p(\chi^2, n - p, \lambda^2 = 0)$; S_b^2/σ^2 is distributed by $p(\chi^2, r, \lambda^2)$ where $\lambda^2 = S_b^2(\xi)/2\sigma^2$, and S_a^2 and S_b^2 are independent. It follows that $u = S_b^2/S_a^2$ is distributed by $p(u, r, n - p, \lambda^2)$ and $w = u/(1 + u) = S_b^2/S_r^2$ is distributed by $p(w, r, n - p, \lambda^2)$.

These results will be established in the following discussion. Let $X = (x_1, \ldots, x_n)$, and $\theta = (\theta_1, \ldots, \theta_p)$ be column vectors and let $C = (c_{ij})$ be the $n \times p$ matrix of the c_{ij}, with rank p. To determine $\hat{\theta}$, we minimize with respect to θ the function

(41) $$f = (X - C\theta)'(X - C\theta) = X'X - \theta'C'X - X'C\theta + \theta'C'C\theta .$$

Differentiating with respect to the θ_j gives the linear equations

(42) $\quad (\partial f/\partial \theta) : (X - C\hat{\theta})'C = 0;$ or $X'C - \theta'C'C = 0.$

Let $D = C'C$. Then $\hat{\theta} = D^{-1}C'X$. Likewise, the estimates of $\hat{\hat{\theta}}$ are determined by minimizing S^2 under H_o. It should be noted that the estimates $\hat{\hat{\theta}}$ satisfy $p - r$ of the equations $D\theta = C'X$. Now for any value of θ,

(43)
$$\begin{aligned}(X - C\theta)'(X - C\theta) &= (X - C\hat{\theta})'(X - C\hat{\theta}) + (\theta - \hat{\theta})'D(\theta - \hat{\theta}) \\ &= S_a^2 + (\theta - \hat{\theta})'D(\theta - \hat{\theta}),\end{aligned}$$

the other terms vanishing because of the least square equations. Hence, for $\theta = \hat{\hat{\theta}}$,

(44)
$$\begin{aligned}(X - C\hat{\hat{\theta}})'(X - C\hat{\hat{\theta}}) &= (X - C\hat{\theta})'(X - C\hat{\theta}) + (\hat{\hat{\theta}} - \hat{\theta})'D(\hat{\hat{\theta}} - \hat{\theta}), \\ S_r^2 &= S_a^2 + (\hat{\hat{\theta}} - \hat{\theta})'D(\hat{\hat{\theta}} - \hat{\theta}), \text{ and} \\ S_b^2 &= (\hat{\hat{\theta}} - \hat{\theta})'D(\hat{\hat{\theta}} - \hat{\theta}) .\end{aligned}$$

Similarly in ω, $\theta_j = \hat{\hat{\theta}}_j = 0; \quad j = p - r + 1, \ldots, p$, and

(45)
$$\begin{aligned}(X - C\theta)'(X - C\theta) &= (X - C\hat{\hat{\theta}})'(X - C\hat{\hat{\theta}}) + (\theta - \hat{\hat{\theta}})'D(\theta - \hat{\hat{\theta}}) \\ &= S_r^2 + (\theta - \hat{\hat{\theta}})'D(\theta - \hat{\hat{\theta}}) .\end{aligned}$$

Combining equations (44) and (45) under H_o, gives

(46)
$$(X - C\theta)'(X - C\theta) = S_a^2 + (\hat{\theta} - \hat{\hat{\theta}})'D(\hat{\theta} - \hat{\hat{\theta}}) + (\theta - \hat{\hat{\theta}})'D(\theta - \hat{\hat{\theta}}).$$

Let $\theta = 0$, then

(47)
$$X'X = S_a^2 + S_b^2 + \hat{\hat{\theta}}'D\hat{\hat{\theta}} .$$

Consider the ranks of these quadratic forms. $X'X$ is of rank n. S_a^2 is the sum of squares of n linear functions of x with p independent linear restrictions of the form $(X - C\hat{\theta})'C = 0$ and thus has rank $r_1 \leq n - p$. Likewise, S_b^2 is a quadratic form in p linear functions of x, subject to $p - r$ independent linear restrictions $D\hat{\theta} = C'X$ and $D\hat{\hat{\theta}} = C'X$ for some $p - r$ equations, and so has rank $r_2 \leq r$. $\hat{\hat{\theta}}'D\hat{\hat{\theta}}$ has rank $r_3 \leq p - r$. Thus $n \leq r_1 + r_2 + r_3 \leq (n - p) + r + (p - r) = n$. Therefore the equality signs hold in each case and by Cochran's Theorem, S_a^2, S_b^2, and $\hat{\hat{\theta}}'D\hat{\hat{\theta}}$ are independently distributed as $\sigma^2 \chi^2$.

To evaluate the distribution parameter λ^2 in the case where the θ are constant parameters, we have $\lambda_1^2 = S_a^2(\xi)/2\sigma^2$ and $\lambda_2^2 = S_b^2(\xi)/2\sigma^2$, where $S_a^2(\xi)$ and $S_b^2(\xi)$ are the values of these forms under the transformation $X = C\theta$. From equation (37) substituting $X = C\theta$ and using $\hat{\theta} = D^{-1}C'X$ we obtain $S_a^2(\xi) = 0$. The value of $S_b^2(\xi)$ must be evaluated in each particular case. However, since $S_b^2 = (\hat{\theta} - \hat{\hat{\theta}})'D(\hat{\theta} - \hat{\hat{\theta}})$, it follows that if $\theta_1 = (\theta_1, \ldots, \theta_{p-r})$ and $\theta_2 = (\theta_{p-r+1}, \ldots, \theta_p)$ and $C = (C_1 \mid C_2)$ where C_1 is n by $p - r$, then

$$(48) \qquad S_b^2(\xi) = \theta_2' \left[(C_2'C_2) - (C_2'C_1)\,(C_1'C_1)^{-1}\,(C_1'C_2)\right] \theta_2$$

From the above it follows in the case where θ are constant parameters, that $S_b^2/S_a^2 = u$ is idstributed by $p(u, r, n - p, \lambda^2)$

given by equation (15) where $\lambda^2 = S_b^2(\xi)/2\sigma^2$ and that $w = u/(1 + u) = S_b^2/S_r^2$ is distributed by $p(w, r, n - p, \lambda^2)$ given by equation (16).

It is evident that the above proofs could have been established with a transformation to the form of equation (6) where z_i is distributed $N(\eta_i, \sigma^2)$, $\Omega : \eta_i = 0, i = p + 1, \ldots, n$; $\omega : \eta_i = 0, i = p - r + 1, \ldots, p$. Under these restrictions,

(49) $$S^2 = \sum_{i=1}^{n} (z_i - \eta_i)^2,\ S_a^2 = \sum_{i=p+1}^{n} z_i^2,\ S_r^2 = \sum_{p-r+1}^{n} z_i^2,$$

$$S_b^2 = \sum_{p-r+1}^{p} z_i^2,\ S_a^2(\eta) = 0,\ S_b^2(\eta) = \sum_{p-r+1}^{p} \eta_i^2.$$

Let us next turn to the case in which the Θ are assumed to be random variables.

Theorem V: Let

(50) $$x_i = \sum_{j=1}^{p} c_{ij}\theta_j + e_i$$

where the θ_j and e_i are independent normal variates with zero means and with variances σ_j^2 and σ^2. Then S_a^2/σ^2 is distributed by $p(\chi^2, n - p, \lambda^2 = 0)$ and under certain conditions S_b^2/σ_b^2 is distributed by $P(\chi^2, r, \lambda^2 = 0)$. If S_b^2 has this distribution, then $v = S_b^2/S_a^2$ is distributed by $p(v, K)$ of equation (36) where $K = \sigma_b^2/\sigma^2$.

Proof: Since S_a^2 and S_b^2 are of ranks $n - p$ and $\underline{r}$ respectively as functions of $\underline{x}$, while the transformation $X = C\theta + E$ is of rank $\underline{n}$, it follows that S_a^2 and S_b^2 are of the same ranks in θ and E. Transformation to the form (6) reduces (50) to the form $z_i = \eta_i + e_i'$, where the e_i' are independent. From (49) it appears that S_a^2 is a function of $e_i'(i = p + 1, \ldots, n)$ and that S_b^2 is a function of $\eta_i + e_i'$, $(i = 1, \ldots, r)$ and hence S_a^2 and S_b^2 are independent.

Now $S_a^2(x) = (X - C\hat{\theta})'(X - C\hat{\theta})$; but $\hat{\theta} = D^{-1}C'X$ and hence $X - C\hat{\theta} = X - CD^{-1}C'X = E - CD^{-1}C'E$. Thus $S_a^2(x) = S_a^2(e)$ and S_a^2 is distributed as $\sigma^2 \chi^2$ with $n - p$ degrees of freedom. Likewise, $S_b^2 = (\hat{\theta} - \hat{\hat{\theta}})'D(\hat{\theta} - \hat{\hat{\theta}})$ under certain conditions is distributed as $\sigma_b^2 \chi^2$ with $\underline{r}$ degrees of freedom. This will be true if the variables of the quadratic form satisfy the distribution conditions of Theorems II or III.

If S_b^2 has the required distribution, then $v = S_b^2/S_a^2$ is distributed as Ku, where σ_b^2 is the variance of elements of S_b^2 and $K = \sigma_b^2/\sigma^2$, $r_1 = r$, $r_2 = n - p$, and v is distributed by $p(v, K)$ given by equation (36).

c. <u>Common tests of the linear hypothesis.</u>

The usual tests of the linear hypothesis are related to the given functions as follows:

(1) Likelihood ratio test.

The likelihood ratio test for the linear hypothesis is

(51) $$L = (S_a{}^2/S_r{}^2)^{n/2}$$

where the critical region is $L < L_o < 1$.

(2) The F-test.

The usual variance ratio test is given by

(52) $$F = (n - p)S_b{}^2/r\, S_a{}^2$$

where the critical region is $F > F_o$.

(3) Fisher's z-test.

(53) $$z = (1/2)\ln F$$

with critical region $z > z_o$.

(4) Tang's E^2 criterion

(54) $$E^2 = S_b{}^2/S_r{}^2 .$$

Since these are all monotonic functions of $S_b{}^2/S_a{}^2$, it can be seen that the tests are equivalent and that the distributions are directly related and can be derived from the distribution of $u = S_b{}^2/S_a{}^2$ or $w = S_b{}^2/S_r{}^2$. The power functions are dependent on the parameters λ^2 and K only.

B. Sequential Analysis

We turn now to a brief review of the principles of sequential analysis. A sequential test of the hypothesis H_o is defined as follows:

Let the m-dimensional sample space be divided into three mutually exclusive parts: R_m^o, R_m^1, R_m^2, $m = 1, \ldots$. At each m, one of the following decisions is made:

(1) If $X = (x_1, \ldots, x_m) \varepsilon R_m^o$, accept H_o.

(2) If $X \varepsilon R_m^1$, reject H_o.

(3) If $X \varepsilon R_m^2$, take another observation.

The selection of R_m^o, R_m^1, R_m^2 for all m, determines a sequential test.

The Operating Characteristic function of a sequential test is defined as follows: If x is distributed by $f(x, \theta)$, then for a given sequential test, the probability of accepting H_o is a function of θ, $L(\theta)$, called the Operating Characteristic or OC function.

In a sequential test, the sample size, n, is a random variable whose distribution is a function of θ. The mean value of this distribution $E_\theta(n)$ is called the expected sample number function or the Average Sample Number function (ASN).

The following principles govern the selection of a sequential test: From a priori considerations, select in the parameter space

(1) A region of acceptance ω_a.

(2) A region of rejection ω_r.

(3) A region of indifference ω_i.

Select upper limits α, β for the probabilities of errors of the first and second kinds, and require that $1 - L(\theta) \leq \alpha$ for $\theta \in \omega_a$, and $L(\theta) \leq \beta$ for $\theta \in \omega_r$. From among the tests satisfying these requirements, select the one having minimum ASN. It has been shown (48) that, for the test of a simple hypothesis $\theta = \theta_o$ against a simple alternative $\theta = \theta_1$, the sequential probability ratio test defined in the following paragraph has the required optimum properties.

The Sequential Probability Ratio Test is defined as follows:

Let $\underline{x}$ be distributed by $f(x, \theta)$. Required, to test the hypothesis H_o: $\theta = \theta_o$ against the alternative H_1: $\theta = \theta_1$. Let $p_{1m} = f(x_1, \ldots, x_m; \theta_1)$ and $p_{om} = f(x_1, \ldots, x_m; \theta_o)$. Let R_m^o be the set $p_{1m}/p_{om} \leq B$, R_m^1 be the set $p_{1m}/p_{om} \geq A$ and R_m^2 the set $B < p_{1m}/p_{om} < A$, where A and B are determined so as to give the desired strength (α, β). For this case, if $p_m = f(x_1, \theta)f(x_2, \theta) \ldots f(x_m, \theta)$, the above test has been shown to be optimum. Tests of more general cases can be constructed so as to satisfy the requirements regarding type one and type two errors but have not been shown to have the desired optimum properties.

It can be shown that $A \leq (1 - \beta)/\alpha$ and $B \geq \beta/(1 - \alpha)$ and that in practice the values $A = (1 - \beta)/\alpha$ and $B = \beta/(1 - \alpha)$ can be used without appreciable increase in α and β and only slight increase in number of observations, $\underline{n}$. When successive observations are independent, the ASN and OC functions can be determined approximately from the formulae:

(55) $$L(\theta) = \frac{A^{h(\theta)} - 1}{A^{h(\theta)} - B^{h(\theta)}},$$

(56) $$E_\theta(n) = \frac{L(\theta)\ln B + [1 - L(\theta)] \ln A}{E_\theta(z)},$$

where $h(\theta)$ is the non-zero solution of

(57) $$\int_{-\infty}^{\infty} [f(x, \theta_1)/f(x, \theta_0)]^{h(\theta)} f(x, \theta)\, dx = 1$$

and

(58) $$z = \ln [f(x, \theta_1)/f(x, \theta_0)] .$$

The method of construction of a general test for a composite hypothesis consists of the following procedure:

Let C be the class of all tests constructed by the following weight function method. Select weight functions $w_a(\theta)$, and $w_r(\theta)$ such that

(59) $$\int_{\omega_a} w_a(\theta)\, d\theta = 1, \quad \int_{\omega_r} w_r(\theta)\, d\theta = 1.$$

Let

(60) $$p_{am} = \int_{\omega_a} f(x_1, \ldots, x_m; \theta)\, w_a(\theta)\, d\theta \quad \text{and}$$

$$p_{1m} = \int_{\omega_r} f(x_1, \ldots, x_m; \theta)\, w_r(\theta)\, d\theta$$

and set up the sequential probability ratio test for p_{1m}/p_{am}. Since

$\alpha(\theta)$ and $\beta(\theta)$, the probabilities of Type I and Type II errors, are functions of θ, require that $\alpha(\theta) \leq \alpha$ for $\theta \in \omega_a$ and $\beta(\theta) \leq \beta$ for $\theta \in \omega_r$. Let $\alpha[A, B, w_a, w_r]$ and $\beta[A, B, w_a, w_r]$ be the maximum of $\alpha(\theta)$ and $\beta(\theta)$ with respect to θ. The optimum test is considered to be that one which minimized these values or some function of them. The process to be used in finding the desired minimum is to consider A and B constant, and find w_a and w_r which give minimum values; then select A and B which make these values equal to the specified α and β.

A general procedure for determining optimum weight functions is not known; in practice the following procedure is suggested: Let

(61) $$w_r(\theta) = v_r(\theta) \text{ on } S_r$$
$$= 0, \text{ otherwise}$$

where S_r is the boundary of ω_r. Let

(62) $$\int_{S_r} v_r(\theta)\, dS_r = 1 \text{ and let}$$
$$p_{1m} = \int_{S_r} f(x_1, \ldots, x_m\, \theta) v_r(\theta)\, dS_r.$$

If the test function p_{1m}/p_{0m} meets the following conditions then it is optimum as defined above:

(1) $\alpha(\theta)$ constant in ω_a.

(2) $\beta(\theta)$ constant on S_r.

(3) $\beta(\theta)$ for θ in $\omega_r \leq \beta(\theta)$ on S_r.

III. SEQUENTIAL TEST OF THE LINEAR HYPOTHESIS

A. Tests Assuming Constant Parameters

Given a general linear hypothesis expressed in the following form:

Let $x_{i\alpha}$,: $i = 1, \ldots, k$; $\alpha = 1, \ldots, m, \ldots$ be distributed normally and independently with mean ξ_i and variance σ^2, and where $\xi_i = \sum_{j=1}^{p} c_{ij}\theta_j$, $p < k$ and C has rank p. Required: to test the hypothesis H_o: $\theta_j = 0$, $j = p - r + 1, \ldots, p$.

1. Test developed by the weight function method.

Using a transformation to the form of equation (6), let $z_{i\alpha} = \Gamma x_{i\alpha}$ where Γ is the same for all i, and such that $z_{i\alpha}$ is distributed $N(\eta_i, \sigma^2)$; $\eta_i = 0$, $i = p + 1, \ldots, k$ and H_o: $\eta_i = 0$, $i = p - r + 1, \ldots, p$. Let $\delta^2 = (1/\sigma^2)\sum_{p-r+1}^{p}\eta_i^2$ be a chosen surface in the space of $(\eta_{p-r+1}, \ldots, \eta_p, \sigma)$. Choose as region of acceptance and rejection the regions

$$\omega_a: \quad \delta^2 = 0; \quad \omega_r: \quad \delta^2 = \delta_o^2 > c. \tag{63}$$

For any fixed sample size $\alpha = 1, \ldots, m$, let

$$S_m^2 = \sum_{i,\alpha}(z_{i\alpha} - \eta_i)^2 = \sum_{\alpha, i=1}^{p}(z_{i\alpha} - \eta_i)^2 + \sum_{\alpha, i=p+1} z_{i\alpha}^2 \tag{64}$$

$$= \sum_{\alpha, i=1}^{p}(z_{i\alpha} - z_{i.})^2 + m\sum_{i=1}^{p}(z_{i.} - \eta_i)^2 + \sum_{\alpha, i=p+1}^{k} z_{i\alpha}^2$$

where $z_{i.} = (1/m)\sum z_{i\alpha}$. Then

(65) $$S_{ma}^2 = \sum_{\alpha, i=1}^{p} (z_{i\alpha} - z_{i.})^2 + \sum_{\alpha, i=p+1}^{k} z_{i\alpha}^{\ 2}.$$

(66) $$S_{mr}^2 = \sum_{\alpha, i=1}^{p} (z_{i\alpha} - z_{i.})^2 + m\sum_{p-r+1}^{p} z_{i.}^{\ 2} + \sum_{\alpha, i=p+1}^{k} z_{i\alpha}^{\ 2}.$$

(67) $$S_{mb}^2 = S_{mr}^2 - S_{ma}^2 = m\sum_{p-r+1}^{p} z_{i.}^{\ 2}$$

are the various minimum values defined in equations (38), (39) and (40). The density function of the $z_{i\alpha}$ is

(68) $$p(z_{11}, \ldots, z_{km}) = (2\pi\sigma^2)^{-mk/2}\, e^{-S_m^2/2\sigma^2}.$$

To construct a sequential test, let

(69) $$p_{1m}/p_{om} = \int_{T_r} p(Z)v_r(\eta, \sigma)\, dT_r \Big/ \int_{\omega_a} p(Z)v_a(\eta, \sigma)\, d\omega_a$$

where T_r is the surface of ω_r and v_r and v_a are the chosen weight functions.

In ω_a:

(70) $$S_m^2 = S_{mr}^2 + m\sum_{1}^{p-r} (z_{i.} - \eta_i)^2.$$

On T_r:

(71) $$S_m^2 = S_{mr}^2 + m\sum_{1}^{p-r} (z_{i.} - \eta_i)^2 + m\,\delta_o^2\sigma^2 - 2\, S_b\sqrt{m}\,\delta_o\sigma\cos\gamma$$

where γ is the angle between the vectors $(z_{p-r+1.}, \ldots, z_{p.})$ and $(\eta_{p-r+1}, \ldots, \eta_p)$. Let

(72) $$v_a(\eta, \sigma) = v_a(\sigma) \cdot \prod_1^{p-r} v(\eta_i),$$

where

(73) $$v_a(\sigma) = 1/c;\ 0 < \sigma \leqslant c,\ \eta_i = 0,\ i = p-r+1, \ldots, p$$
$$= 0, \quad \text{otherwise},$$

and

(74) $$v(\eta_i) = 1/2a;\ -a \leqslant \eta_i \leqslant a$$
$$= 0;\ \text{otherwise}.$$

Let

(75) $$v_r(\eta, \sigma) = v_r(\sigma) \prod_1^{p-r} v(\eta_i), \text{ where}$$

(76) $$v_r(\sigma) = \Gamma(r/2)/\, c\sqrt{(1+\delta_o^2)}\, 2\pi^{r/2} \delta_o^{\,r-1} \sigma^{r-1};$$
$$0 < \sigma \leqslant c;\quad \delta^2 = \delta_o^2$$
$$= 0;\ \text{otherwise}.$$

Substituting in equation (69) and letting $\underline{a}$ and $\underline{c}$ approach infinity, we obtain

(77) $$p_{1m}/p_{om} = e^{-m\delta_o^2/2} \int_{T_r'} \sigma^{-mk}$$
$$\exp\left\{(-1/2\sigma^2)\left[S_r^2 - 2\sqrt{m}\,\sigma\,\delta_o\, S_b \cos\gamma\right]\right\} v_r'(\sigma)\, dT_r'$$
$$\prod_{i=1}^{p-r} \int_{-\infty}^{\infty} \exp\left\{(-1/2\sigma^2) m (z_{i.} - \eta_i)^2\right\} d\eta_i$$

$$\div \int_0^{\infty} \sigma^{-mk} \exp\left\{- S_r^2/2\sigma^2\right\} d\sigma$$

$$\prod_{i=1}^{p-r} \int_{-\infty}^{\infty} \exp\left\{(-1/2\sigma^2) m (x_{i.} - \eta_i)^2\right\} d\eta_i$$

where T'_r is the surface of $\sum \eta_i^2 = \delta_o^2 \sigma^{-2}$ and $v_r'(\sigma) = cv_r(\sigma)$.

The differential element dT_r' can be expressed as follows: Let dt be a segment of an element of the hyper-cone T_r' and let α be the angle between this element and the vector $(0, 0, \ldots, 0, \sigma)$. Then $\cos\alpha = 1/\sqrt{(1+\delta_o^2)}$, $dt = \sqrt{(1+\delta_o^2)}\, d\sigma$, and $dT_r' = dt\, dS$, where dS is a surface element of the hyper-sphere $\sum \eta_i^2 = \delta_o^2 \sigma^2$, σ^2 constant. Since the integrand varies on this sphere only as the central angle γ, it follows that dS can be written as $\delta_o\, \sigma\, d\varphi\, S'$, where S' is the area of a hyper-sphere in $r - 1$ dimensions with radius $\delta_o\,\sigma \sin\varphi$ and φ is any central angle. Thus

$$dT_r' = \sqrt{(1+\delta_o^2)}\, d\sigma\, (\delta_o\, \sigma\, d\varphi)\, 2\pi^{(r-1)/2} \delta_o^{\,r-2}\, \sigma^{r-2} \sin^{r-2}\varphi / \Gamma\left[(r-1)/2\right]. \tag{78}$$

A similar result can be achieved using a polar transformation. On making the above substitution,

$$p_{1m}/p_{om} = e^{-m\,\delta_o^2/2}\, \Gamma(r/2)/\sqrt{\pi}\, \Gamma\left[(r-1)/2\right] \int_0^{\infty}\int_0^{\pi} \sigma^{-mk+p-r} \exp\left\{(-1/2\sigma^2)\left[S_r^2 - 2S_b\, \delta_o \sqrt{m}\, \sigma \cos\varphi\right]\right\} \sin^{r-2}\varphi\, d\varphi\, d\sigma \div \int_0^{\infty} \sigma^{-mk+p-r} \exp\left\{- S_r^2/2\sigma^2\right\} d\sigma. \tag{79}$$

It can be shown that p_{1m}/p_{om} is a homogeneous function of zero degree in $z_{i\alpha}$ and therefore is unaltered by the substitution of $z_{i\alpha}/S_r^2$ for $z_{i\alpha}$, giving

(80) $$p_{1m}/p_{om} = e^{-m\delta_o^2/2} B^{-1}\left[(r-1)/2, 1/2\right]$$

$$\int_0^\infty\int_0^\pi \sigma^{-mk+p-r}\exp\left\{-1/2\sigma^2 + w\,\delta_o\sqrt{m}\cos\varphi/\sigma\right\}$$

$$\sin^{r-2}\varphi\, d\varphi\, d\sigma \div \int_0^\infty \sigma^{-mk+p-r}\exp\left\{-1/2\sigma^2\right\} d\sigma .$$

Now

(81) $$\int_0^\pi e^{\delta_o w\sqrt{m}\cos\varphi/\sigma}\sin^{r-2}\varphi\, d\varphi =$$

$$\sum_{j=0}^\infty (m\,\delta_o^2)^j w^{2j}/\sigma^{2j}\,(2j! \int_0^\pi \sin^{r-2}\varphi\cos^{2j}\varphi\, d\varphi$$

$$= \sum_{j=0}^\infty w^j(\delta^2{}_o m)^j/\sigma^{2j}\,(2j)!\, B\left[(r-1)/2, j+1/2\right]$$

$$= \sqrt{\pi}\,\Gamma\left[(r-1)/2\right]\sum_{j=0}^\infty\left[(m\,\lambda_o^2)^j w^j/j!\,\sigma^{2j}\right]\cdot\left[1/\Gamma(r/2+j)\,2^j\right]$$

where $\lambda_o^2 = \delta_o^2/2$.

(82) $$p_{1m}/p_{om} = e^{-m\lambda_o^2}\,\Gamma(r/2)\sum_{j=0}^\infty\left[(m\,\lambda_o^2)^j w^j/j!\,\Gamma(r/2+j)2^j\right]$$

$$\int_0^\infty \sigma^{-mk+p-r-2j} e^{-1/2\sigma^2} d\sigma \div \int_0^\infty \sigma^{-mk+p-r} e^{-1/2\sigma^2} d\sigma$$

$$= B[r/2, (mk-p-1)/2]\ e^{-m\lambda_o^2} \sum_{j=0}^{\infty} (m\lambda_o^2)^j w^j \bar{B}^{1}[(mk-p-1)/2, r/2+j]$$

(83) $$p_{1m}/p_{om} = e^{-m\lambda_o^2} F[(mk-p+r-1)/2, r/2, m\lambda_o^2 w]$$

2. Test by the direct use of the w-distribution

The fact that the weight function approach results in a test function dependent on the distribution of w, suggests the possibility of the direct use of this distribution in construction of a test function.

As in the preceding section, we have

(84) $$s_m^2 = \sum_{i,\alpha} (x_{i\alpha} - \sum_{j=1}^{p} c_{ij}\theta_j)^2$$

$$= \sum (x_{i\alpha} - \bar{x}_i)^2 + m\sum_i (\bar{x}_i - \sum_j c_{ij}\theta_j)^2$$

and thus $\hat{\theta} = S^{-1}C'\bar{X}$, and $\hat{\hat{\theta}} = S'^{-1}_{p-r}\, C'_{p-r}\, \bar{X}$ are the same functions of $\bar{X}_m$ for any value of m. Further

(85) $$s_{ma}^2 = \sum_{i,\alpha} (x_{i\alpha} - \bar{x}_i)^2 + m\sum_i (\bar{x}_i - \sum_j c_{ij}\hat{\theta}_j)^2$$

$$s_{mr}^2 = \sum_{i,\alpha} (x_{i\alpha} - \bar{x}_i)^2 + m\sum_i (\bar{x}_i - \sum_j c_{ij}\hat{\hat{\theta}}_j)^2$$

$$s_{mb}^2 = m\sum_i (\bar{x}_i - \sum_j c_{ij}\hat{\hat{\theta}}_j)^2 - m\sum_i (\bar{x}_i - \sum_j c_{ij}\hat{\theta}_j)^2.$$

Or in matrix notation:

$$(86) \qquad S_m^2 = Q + m(\bar{X} - C\theta)'(\bar{X} - C\theta)$$

$$S_{ma}^2 = Q + m(\bar{X} - C\hat{\theta})'(\bar{X} - C\hat{\theta})$$

$$S_{mr}^2 = Q + m(\bar{X} - C\hat{\hat{\theta}})'(\bar{X} - C\hat{\hat{\theta}})$$

where $Q = \sum_{i,\alpha} (x_{i\alpha} - \bar{\bar{x}}_i)^2$. Hence

$$(87) \qquad S_{mb}^2 = m(\hat{\theta} - \hat{\hat{\theta}})'S(\hat{\theta} - \hat{\hat{\theta}})$$

and from Theorem II the parameter

$$(88) \qquad \lambda_m^2 = S_{mb}^2(\theta)2\,\sigma^2 = mS_{b1}^2(\theta)/2\sigma^2 = m\,\lambda_1^2.$$

It follows from Theorem IV that $w_m = S_{mb}^2/S_{mr}^2$ is distributed by

$$(89) \qquad p(w_m) = e^{-m\,\lambda^2}\,B^{-1}(r_1/2,\ r_2/2)\,w^{r_1/2 - 1}\,(1 - w)^{r_2/2 - 1}$$

$$F\left[(r_1 + r_2)/2,\ r_2/2,\ m\,\lambda^2 w_m\right]$$

where $r_1 = r$ and $r_2 = mk - p$. The sequential probability ratio test for H_0: $\lambda^2 = \lambda_0^2 \geqslant 0$ versus H_1: $\lambda^2 = \lambda_1^2 > \lambda_0^2$ gives

$$(90) \qquad p_{1m}/p_{0m} = e^{-m(\lambda_1^2 - \lambda_0^2)}\,F\left[(r_1 + r_2)/2,\ r_2/2,\ m\,\lambda_1^2\,w_m\right]$$

$$\div\ F\left[(r_1 + r_2)/2,\ r_2/2,\ m\,\lambda_0^2\,w_m\right].$$

Verification needed? (i.e. use of theorem on sufficiency etc.?)

If $\lambda_o^2 = 0$, this test function reduces to

$$(91) \qquad p_{1m}/p_{om} = e^{-m\lambda^2} F\left[(mk-p+r)/2,\ r/2,\ m\lambda^2 w_m\right]$$

and is thus nearly equivalent to the test function of equation (83).

This provides a proper sequential ratio test of the simple hypothesis $\lambda^2 = \lambda_o^2$ against the single alternative $\lambda^2 = \lambda_1^2$, for which the probability of Type I error $P_I = \alpha$ and probability of Type II error $P_{II} = \beta$. There are several difficulties in the application of these two tests. First, since $\alpha(\lambda^2)$ and $\beta(\lambda^2)$, the probabilities of Type I and Type II errors, depend on the joint distribution of a sequence of values of p_{1m}/p_{om}, it has not been established that the test ratio meets the required conditions regarding these probabilities. Second, since p_{1m}/p_{om} is not a product of independent ratios, a number of methods and theorems of sequential analysis based on this property are inapplicable. In particular, the methods of evaluation of the ASN and OC functions given in equations (53) to (55) cannot be used. Finally, the values w_m must be calculated for each $\underline{m}$, using all previous observations. Also the parameter $r_2 = mk - p$ of the hypergeometric function is continually increasing and necessitates the evaluation of this function over a large range of this parameter. Some of these difficulties can be eliminated when the following construction is appropriate.

3. <u>Test by the ratio of independent w-distributions.</u>

Let the linear hypothesis be the same as before. For each sample $\alpha = q$, calculate $w = S_{qb}^2/S_{qr}^2$ based on this sample <u>only</u>. Thus we secure a set $w_1, \ldots, w_m, \ldots$ of independent values with a common distribution. Construct a test function

$$(92)\qquad p_{1m}/p_{om} = p(w_1, \ldots, w_m;\ \lambda_1^2)/p(w_1, \ldots, w_m;\ \lambda_o^2)$$

$$= \prod_{q=1}^{m} p(w_q,\ \lambda_1^2)/p(w_q,\ \lambda_o^2).$$

For any particular value q of α we have

$$(93)\qquad S_q^{\ 2} = \sum_i (x_{iq} - \sum_{j=1} c_{ij}\theta_j)^2$$

$$S_{qa}^2 = \sum_i (x_{iq} - \sum_j c_{ij}\hat{\theta}_j)^2$$

$$S_{qr}^2 = \sum_i (x_{iq} - \sum_j c_{ij}\hat{\hat{\theta}}_j)^2$$

$$S_{qb}^2 = S_{qr}^2 - S_{qa}^2$$

and $w_q = S_{bq}^2/S_{rq}^2$ is distributed by $p(w,\ r_1,\ r_2,\ \lambda^2)$ where $r_1 = r$, $r_2 = k - p$, $\lambda^2 = S_b^2(\theta)/2\sigma^2$. To test $\lambda^2 = \lambda_o^2 \geqslant 0$ against $\lambda^2 = \lambda_1^2 > \lambda_o^2$, let

$$(94)\qquad p_{1m}/p_{om} = \prod_{q=1}^{m} p(w_q,\ \lambda_1^2)/p(w_q,\ \lambda_o^2)$$

$$= e^{-m(\lambda_1^2 - \lambda_o^2)} \prod_{q=1}^{m} \Big\{ F\left[(r_1 + r_2)/2, r_2/2, \lambda_1^2\ w_q\right]$$

$$\div F\left[(r_1 + r_2)/2,\ r_2/2,\ \lambda_o^2\ w_q\right] \Big\}$$

or in logarithmic form:

$$\text{(95)} \qquad \ln(p_{1m}/p_{om}) = -m(\lambda_1^2 - \lambda_o^2) + \sum_{q=1}^{m}\left\{\ln F\left[(r_1 + r_2)/2,\ r_2/2,\ \lambda_1^2 w_q\right] - \ln F\left[(r_1 + r_2)/2,\ r_2/2,\ \lambda_o^2\ w_q\right]\right\} .$$

This test provides a test of the simple hypothesis $\lambda^2 = \lambda_o^2 \geqslant 0$ against the simple alternative $\lambda^2 = \lambda_1^2 > \lambda_o^2$ with probabilities of Type I error and Type II error of α and β respectively. It can also be shown that the OC function $L(\lambda^2)$ is a decreasing function of λ^2. To show this we proceed as follows.

Let x be distributed by $f(x, \theta)$, and let the hypothesis $\theta = \theta_o$ be tested against the alternative $\theta = \theta_1 > \theta_o$ by the sequential ratio

$$\text{(96)} \qquad p_{1m}/p_{om} = \frac{f(x_1, \theta_1) \cdot \cdot \cdot \cdot \cdot \cdot \cdot f(x_m, \theta_1)}{f(x_1, \theta_o) \cdot \cdot \cdot \cdot \cdot \cdot \cdot f(x_m, \theta_o)} .$$

Let (a) $f(x, \theta_1)/f(x, \theta_o)$ be a monotonic increasing function of x for any $\theta_1 > \theta_o$ and (b) $\text{Prob}(x > c \mid \theta)$ be an increasing function of θ for any c. Then it can be shown that $L(\theta)$ is a decreasing function of θ. From equation (55) it can be shown that $L(\theta)$ is an increasing function of $\underline{h}$. The relation between h and θ, as defined by equation (57), can be written

$$\text{(97)} \qquad g(h, \theta) = E\left[f(x, \theta_1)/f(x, \theta_o)\right]^h = \int_{-\infty}^{\infty}\left[f(x, \theta_1)/f(x, \theta_o)\right]^h f(x, \theta)\, dx = 1.$$

It follows from the first two conditions that $g(h, \theta)$ is an increasing function of θ for $h > 0$ and a decreasing function for $h < 0$. Also it is known((46) pp.158 - 159) that $\partial g/\partial h \big|_{g = 1}$ is positive when $h > 0$ and negative for $h < 0$. Hence $\left.\frac{dh}{d\theta}\right|_{g = 1} = - \left.\frac{g_\theta(h, \theta)}{g_h(h, \theta)}\right|_{g = 1} < 0$ and h is a decreasing function of θ. Therefore $L(\theta)$ is a decreasing function of θ.

For the test based on independent w-distributions, it can be shown that $F(a, b, \lambda_1^2 w)/F(a, b, \lambda_o^2 w)$ is an increasing function of $\underline{w}$. Also it has been shown in the discussion of the w-distribution that $\text{Prob}(w > w_o \mid \lambda^2)$ is an increasing function of λ^2. Therefore it is apparent that this test satisfies the requirements stated above and that $L(\lambda^2)$ is a decreasing function of λ^2.

It appears that this test has several advantages from the standpoint of application but that it is probably less powerful than the first test functions suggested.

4. Example of the construction of a test function.

As an example of the application of the above methods to the construction of various test functions let us consider the construction of a sequential t-test. Given x_α ; $\alpha = 1, \ldots, m, \ldots,$ x_α distributed $N(\theta, \sigma^2)$; required to test the hypothesis H_o: $\theta = 0$.

a. <u>Test based on the weighted integral</u>

Here we have the construction only for $\lambda_o^2 = 0$.

(98) $$S_{mr}^2 = \sum x_\alpha^2, \quad S_{mb}^2 = m\bar{x}^2, \quad w_m = S_{mb}^2/S_{mr}^2,$$

$$S_{mb}^2(\theta) = m\theta^2, \quad \lambda_m^2 = m\theta^2/2\sigma^2,$$

(99) $$P_{1m}/P_{om} = e^{-m\lambda^2} F\left[(m-1)/2,\ 1/2,\ m\lambda^2 w_m\right].$$

b. <u>Wald's t-test</u>

Wald ((46) pp. 205) gives the test function

(100) $$P_{1m}/P_{om} = (1/2)\int_0^\infty \sigma^{-m}\Big[e^{-(1/2\sigma^2)\sum(x_\alpha - \theta_o - \delta\sigma)^2} + e^{-(1/2\sigma^2)\sum(x_\alpha - \theta_o + \delta\sigma)^2}\Big]\, d\sigma \div \int_0^\infty \sigma^{-m} e^{-(1/2\sigma^2)\sum(x_\alpha - \theta_o)^2}\, d\sigma$$

where $\delta = |\theta - \theta_o|/\sigma$ and thus $\lambda^2 = \delta^2/2$. To correspond with our present hypothesis, let $\theta_o = 0$. Then

(101) $$P_{1m}/P_{om} = \frac{e^{-m\lambda^2}}{2}\int_0^\infty \sigma^{-m} e^{-S_r^2/2\sigma^2}\left[e^{-m\lambda\sqrt{2}\,\bar{x}/\sigma} + e^{-m\lambda\sqrt{2}\,\bar{x}/\sigma}\right] \cdot d\sigma \div \int_0^\infty \sigma^{-m} e^{-S_r^2/2\sigma^2}\, d\sigma$$

where $S_r^2 = \sum(x_\alpha^2)$. Since this function is homogeneous of zero degree in x, we can make the transformation $x_\alpha = x_\alpha/S_r$ giving

(102) $$P_{1m}/P_{om} = \frac{e^{-m\lambda^2}}{2^{(m-1)/2}\Gamma[(m-1)/2]}\int_0^\infty \sigma^{-m} e^{-1/2\sigma^2}\left[e^{-\lambda\sqrt{2mw}/\sigma} + e^{+\lambda\sqrt{2mw}/\sigma}\right] d\sigma$$

where $w = m\bar{x}^2/S_r^2$. Expanding in power series and integrating,

(103) $$p_{1m}/p_{om} = e^{-m\lambda^2} F\left[(m-1)/2,\ 1/2,\ m\lambda^2 w\right]$$

and therefore these two tests are equivalent.

c. <u>Test based on the w_m distribution.</u>

Using the sums of squares of equations (98) and the test ratio of equation (90) we have

(104) $$p_{1m}/p_{om} = e^{-m(\lambda_1^2 - \lambda_o^2)} \frac{F\left[m/2,\ 1/2,\ m\lambda_1^2 w_m\right]}{F\left[m/2,\ 1/2,\ m\lambda_o^2 w_m\right]} .$$

d. <u>Test based on pairs of sample values</u>

Let the sample be taken by drawing pairs of values and construct a test function on the basis of the w distribution for these pairs. Let the sample values be designated by

(105) $$x_{i\alpha} : i = 1, 2 ; \quad \alpha = 1, \ldots, m, \ldots$$

and let $E(x_{i\alpha}) = \theta$. For any pair, say x_{1q}, x_{2q} we have

(106) $$S_q^2 = \sum_i (x_{iq} - \theta)^2$$

$$S_{qa}^2 = \sum_i (x_{iq} - \bar{x}_q)^2, \quad r_2 = 1;$$

$$S_{qr}^2 = \sum_i x_{iq}^2, \quad S_{qb}^2 = 2\bar{x}_q^2, \quad r_1 = 1;$$

$$\lambda^2 = \theta^2/\sigma^2; \quad w_q = S_{qb}^2/S_{qr}^2.$$

$$(107) \qquad p_{1m}/p_{om} = e^{-m(\lambda_1^2 - \lambda_o^2)} \prod_{q=1}^{m} F(1, 1/2, \lambda_1^2 w_q)/F(1, 1/2, \lambda_o^2 w_q) \quad .$$

In applying this test it is probably convenient to use the function

$$(108) \qquad L_m = \ln(p_{1m}/p_{om}) = -m(\lambda_1^2 - \lambda_o^2) + \sum_{q=1}^{m} \ln \frac{F(1, 1/2, \lambda_1^2 w_q)}{F(1, 1/2, \lambda_o^2 w_q)}$$

and to test

$$(109) \qquad \sum \ln \frac{F(1, 1/2, \lambda_1^2 w_q)}{F(1, 1/2, \lambda_o^2 w_q)} \quad \text{against}$$

$$(110) \qquad A' = m(\lambda_1^2 - \lambda_o^2) + \ln A \text{ and}$$

$$B' = m(\lambda_1^2 - \lambda_o^2) + \ln B.$$

To calculate w_q, use

$$(111) \qquad w_q = (x_{1q} + x_{2q})^2/2(x_{1q}^2 + x_{2q}^2).$$

$F(1, 1/2, \lambda^2 w_q)$ can be evaluated from a graph or table of this function.

B. Tests When The Parameters Are Random Variables

Let us consider possible sequential tests of the linear hypothesis when the parameters are assumed to be random variables.

Given: $x_{i\alpha} = \sum_{j=1}^{p} c_{ij}\theta_j + e_{i\alpha}$; $i = 1, \ldots, k$; $\alpha = 1, \ldots, m, \ldots$, where θ_j are normally and independently distributed by $N(0, \sigma_j^2)$, $e_{i\alpha}$ are $N(0, \sigma^2)$ and θ and $\underline{e}$ are independent. The hypothesis to be tested is that $\sigma_j^2 = 0$, $j = p - r + 1, \ldots, p$.

1. Test based on the ratio of two v_m distributions.

Let

$$(112)\qquad S_m^2 = \sum_{i,\alpha} (x_{i\alpha} - \sum_j c_{ij}\theta_j)^2 = \sum_{i,\alpha} (x_{i\alpha} - x_{i.})^2 + m \sum_i (x_{i.} - \sum_j c_{ij}\theta_j)^2$$

and let

$$(113)\qquad Q_m = \sum_{i,\alpha} (x_{i\alpha} - x_{i.})^2; \quad x_{i.} = (1/m) \sum_{\alpha=1}^{m} x_{i\alpha} \cdot$$

Changing to matrix notation,

$$(114)\qquad \begin{aligned} S_m^2 &= Q_m + m(X. - C\theta)'(X. - C\theta) \\ S_{ma}^2 &= Q_m + m(X. - C\hat{\theta})'(X. - C\hat{\theta}) \\ S_{mr}^2 &= Q_m + m(X. - C\hat{\hat{\theta}})'(X. - C\hat{\hat{\theta}}) \\ S_{mb}^2 &= m(\hat{\theta} - \hat{\hat{\theta}})'D(\hat{\theta} - \hat{\hat{\theta}}) \text{ and} \\ \hat{\theta}_m &= D^{-1} C'X., \quad \hat{\hat{\theta}}_m = D_{p-r}^{-1} C'_{p-r} X. \end{aligned}$$

where C_{p-r} is the matrix $\begin{pmatrix} c_{11} & \cdots & c_{p-r} \\ \cdot & \cdots & \cdot \\ c_{k1} & \cdots & c_{k,p-r} \end{pmatrix}$

and D_{p-r}^{-1} is the reciprocal of $C'_{p-r}C_{p-r} = D_{p-r}$.

It follows from Theorem V that S_{ma}^2 and S_{mb}^2 are independent quadratic forms with ranks $mk - p$ and $\underline{r}$ respectively, and that S_{ma}^2 is distributed as $\sigma^2\chi^2$ with $mk - p$ degrees of freedom. Also, under certain conditions S_{mb}^2 will be distributed as $m\sigma_b^2\chi^2$ with $\underline{r}$ degrees of freedom. In particular, this will be true when $x_{\cdot}$ are independent and have equal variance.

Thus $\sigma^2 S_{mb}^2/m\sigma_b^2 S_{ma}^2 = v_m/mK$ where $K = \sigma_b^2/\sigma^2$ is distributed as $p(u, r_1, r_2)$ and so

$$(115) \qquad p(v_m, K) = (mK)^{r_2/2} B^{-1}(r_1/2, r_2/2) \, v^{r_1/2 - 1}(mK + v)^{-(r_1 + r_2)/2}$$

where $r_1 = r$ and $r_2 = mk - p$. Where the exact form of K is to be determined for each particular problem. A test of the hypothesis $K = K_o \geqslant 1$ against $K = K_1 > K_o$ can be formed by taking the probability ratio

$$(116) \qquad p_{1m}/p_{om} = p(v_m, K_1)/p(v_m, K_o) = (K_1/K_o)^{r_2/2} \left[(mK_o + v_m)/mK_1 + v_m)\right]^{(r_1 + r_2)/2}.$$

Since this test ratio is not a product of m independent ratios, a part of the theory of sequential tests is again inapplicable. In particular, it has not been shown that the required conditions regarding $\alpha(K)$ and $\beta(K)$ are satisfied. As in the case of constant coefficients, an alternative test can be constructed using independent distributions.

2. Test by the use of independent v-distributions.

For any particular value of α, say q, we have

$$(117)\qquad S_{qa}^2 = \sum_i \left(x_{iq} - \sum_{j=1}^{p} c_{ij}\hat{\theta}_j\right)^2$$

which is distributed as $\sigma^2\chi^2$ with k - p degrees of freedom. Also $S_{qb}^2 = (\hat{\theta} - \hat{\hat{\theta}})'S(\hat{\theta} - \hat{\hat{\theta}})$ under certain conditions is distributed as $\sigma_b^2\chi^2$ with r degrees of freedom. Then

$$(118)\qquad u = S_{qb}^2 \cdot \sigma^2/S_{qa}^2 \cdot \sigma_b^2 = v_q/K$$

and

$$(119)\qquad p(v_q, K) = K^{r_2/2}\, B^{-1}(r_1/2, r_2/2)v^{r_1/2 - 1} \cdot (K + v)^{-(r_1+ r_2)/2}$$

To test $K = K_o \geqslant 1$ against the alternative $K = K_1 > K_o$ use the ratio

$$(120)\qquad p_{1m}/p_{om} = \prod_{q=1}^{m} \left[p(v_q, K_1)/p(v_q, K_o)\right]$$

$$= (K_1/K_o)^{mr_2/2} \prod_{q=1}^{m} \left[(K_o + v_q)/(K_1 + v_q)\right]^{(r_1 + r_2)/2} .$$

This test ratio provides a test of the hypothesis $K = K_o$ against $K = K_1$ with probabilities of Type I and Type II errors of α and β respectively. Also the OC function is a decreasing function of K. For

(121) $$p(v, K_1)/p(v, K_o) = (K_1/K_o)^{r_2/2} (K_o + v)/K_1 + v) \quad ^{(r_1 + r_2)/2}$$

is an increasing function of v. Also

(122) $$\operatorname{Prob}(v < C \mid K) = B^{-1}(r_1/2, r_2/2) \int_0^{C/K} u^{r_1/2 - 1}(1 + u)^{-(r_1 + r_2)/2}\, du,$$

(123) $$d\operatorname{Prob}(v < C \mid K)/dk = -B^{-1}(r_1/2, r_2/2)(C/K)^{r_1/2 - 1} (1 + C/K)^{-(r_1 + r_2)/2}\, C/K^2 < 0$$

and thus $\operatorname{Prob}(v > C \mid K)$ is an increasing function of K. Therefore this test meets the conditions required following equation (96) and L(K) is a decreasing function of K. The equations for evaluating the ASN and OC functions can be set up in terms of equations (55) and (58).

3. <u>Example of the construction of test functions</u>

As an example of the construction of such tests, consider a randomized block experiment, with <u>t</u> treatments, in <u>b</u> blocks with block-treatment interaction assumed to be zero. Required to test the hypothesis that the variance of treatment effects, σ_t^2 is zero. Let $x_{ij\alpha}$ be the α observation on the i-th treatment in the j-th block, where $(i = 1, \ldots, t)$,

$(j = 1, \ldots, b)$, $\alpha = 1, \ldots, m$, ... Assume the linear hypothesis

(125) $$x_{ij\alpha} = m + t_i + b_j + e_{ij\alpha}, \quad \sum t_i = 0.$$

a. Test based on v_m.

(126) $$S_{ma}^2 = \sum_{i,j,\alpha} (x_{ij\alpha} - x_{ij.})^2 + m \sum_{i,j} (x_{ij.} - x_{i..} - x_{.j.} + x_{...})^2$$

$$x_{ij.} = m + t_i + b_j + e_{ij.}$$

$$x_{i..} = m + t_i + b_. + e_{i..}$$

$$x_{.j.} = m + b_j + e_{.j.}$$

$$x_{...} = m + b_. + e_{...}$$

(127) $$S_{ma}^2 = \sum_{i,j,\alpha} (e_{ij} - e_{ij.})^2 + m \sum_{i,j} (e_{ij.} - e_{i..} - e_{.j.} + e_{...})^2 .$$

Hence S_{ma}^2 is distributed as $\sigma^2\chi^2$ with $(m - 1)bt + (b - 1)(t - 1)$ degrees of freedom

(128) $$S_{mb}^2 = mb \sum_i (x_{i..} - x_{...})^2 = mb \sum_i (t_i + e_{i..} - e_{...})^2$$

$$E\, S_{mb}^2 = mb(t - 1)(\sigma_t^2 + \sigma^2/mb) .$$

Hence S_{mb}^2 is distributed as $(\sigma^2 + mb\,\sigma_t^2)\chi^2$ with $(t - 1)$ degrees of freedom. It follows that $v_m = S_{mb}^2/S_{ma}^2$ is distributed by $p(v; K)$, where $K = 1 + mC$ and $C = b\,\sigma_t^2/\sigma^2$. To test the simple hypothesis $C = C_o \geqslant b$ against the alternative $C = C_1 > C_o$ we form the test ratio

$$(129) \qquad p_{1m}/p_{om} = \left[(1 + mC_1)/(1 + mC_o)\right]^{r_2/2} \left[(1 + mC_1 + v_m)/(1 + mC_o + v_m)\right]^{-(r_1 + r_2)/2}$$

where $r_1 = t - 1$, $r_2 = (m - 1)bt + (b - 1)(t - 1)$.

b. <u>Test based on independent v_i.</u>

For a given α say $\alpha = q$

$$(130) \qquad S_{qa}^2 = \sum_{i,j} (x_{ijq} - x_{.jq} - x_{i.q} + x_{..q})^2$$

is distributed as $\sigma^2\chi^2$ with $(t - 1)(b - 1)$ degrees of freedom.

$$(131) \qquad S_{qb}^2 = b \sum_{i} (x_{i.q} - x_{..q})^2$$

is distributed as $(\sigma^2 + b\,\sigma_t^2)\chi^2$ with $t - 1$ degrees of freedom. Hence $v_q = S_{qb}^2/S_{qa}^2$ is distributed by $p(v, K)$ where $K = (\sigma^2 + b\,\sigma_t^2)/\sigma^2 = 1 + C$, $C = b\,\sigma_t^2/\sigma^2$. To test $C = C_o \geqslant b$ against the alternative $C = C_1 > C_o$, form the test ratio

$$(132) \qquad p_{1m}/p_{om} = \prod_{q=1}^{m} \left[p(v_q;\ c_1)/p(v_q;\ c_o) \right]$$

$$= \left[(1 + c_1)/(1 + c_o) \right]^{mr_2/2}$$

$$\prod_{q=1}^{m} \left[(1 + c_o + v_q)/(1 + c_o + v_q) \right]^{(r_1 + r_2)/2}.$$

IV. SUMMARY

The various forms of the linear hypothesis and the distributions of the test functions for the non-central case have been discussed. A short review of the principles of sequential analysis was followed by application to the problem of sequential tests for the linear hypothesis.

For the case in which the parameters are assumed to be unknown constants, the following tests were developed:

(1) A test of the hypothesis $\lambda^2 = 0$ against the alternative $\lambda^2 = \lambda_o^2 > 0$ using the weight function method of construction. This is a generalization of Wald's sequential t-test.

(2) A test of the hypothesis $\lambda^2 = \lambda_o^2 \geqslant 0$ against the alternative $\lambda^2 = \lambda_1^2 > \lambda_o^2$, based on the direct use of the w-distribution. This is nearly equivalent to the preceding test but is more general.

(3) A test of the hypothesis $\lambda^2 = \lambda_o^2 \geqslant 0$ against the alternative $\lambda^2 = \lambda_1^2 > \lambda_o^2$ based on successive independent variance ratios. It was shown that this test is also a proper test

of $\lambda^2 \leqslant \lambda_o^2$ against $\lambda^2 \geqslant \lambda_1^2$.

For the case in which the parameters are assumed to be random variables, the following tests were developed:

(1) A test of the hypothesis $K = K_o \geqslant 1$ against the alternative $K = K_1 > K_o$, based on the direct use of the v-distribution.

(2) A test of the hypothesis $K = K_o \geqslant 1$ against the alternative $K = K_1 > K_o$, based on successive independent variance ratios. It was shown that this test is also a proper test of $K \leqslant K_o$ against $K \geqslant K_1$.

Examples were given of the construction of a sequential t-test and of a test for a randomized block experiment.

V. BIBLIOGRAPHY

(1) Anderson, T. W. The non-central Wishart distribution and certain problems of multivariate statistics. Ann. Math. Stats. 17: 409 - 431. 1946.

(2) Barnard, G. A. Sequential tests in industrial statistics. Suppl. J. R. S. S. 8: 1 - 26. 1946.

(3) Blackwell, D. On an equation of Wald. Ann. Math. Stats. 17: 84 - 86. 1946.

(4) Blackwell, D. Conditional expectation and unbiased sequential estimation. Ann. Math. Stats. 18: 105 - 110. 1947 .

(5) ———— and Girshick, M. A. A lower bound for the variance of some unbiased sequential estimates. Ann. Math. Stats. 18: 277 - 279. 1947.

(6) Burman, J. P. Sequential sampling formulae for a binomial population. Suppl. J. R. S. S. 8: 98 - 103. 1946 .

(7) Cochran, W. G. The distribution of quadratic forms in a normal system, with applications to the analysis of covariance. Proc. Camb. Phil. Soc., 30: 178 - 191. 1934.

(8) Craig, A. T. On the independence of certain estimates of variance. Ann. Math. Stats. 9: 48 - 55. 1938.

(9) ———— Note on the independence of certain quadratic forms. Ann. Math. Stats. 14: 195 - 196. 1943.

(10) ———— Bilinear forms in normally correlated variables. Ann.Math. Stats. 18: 565 - 575. 1947.

(11) Craig, C. C. Note on the distribution of non-central $\underline{t}$ with an application. Ann. Math. Stats. 12: 224 - 227. 1941 .

(12) Fisher, R. A. The distribution of the partial correlation coefficient. Metron 3: 329 - 337. 1924 .

(13) Fisher, R. A. Statistical methods for research workers. 9th ed. Edinburgh, Oliver and Boyd. 1944.

(14) ______ The general sampling distribution of the multiple correlation coefficient. Proc. Roy. Soc., A, 121: 654 - 659. 1928.

(15) Girshick, M. A. Contributions to the theory of sequential analysis I. Ann. Math. Stats. 17: 123 - 143. 1946.

(16) ______ Contributions to the theory of sequential analysis. II, III. Ann. Math. Stats. 17: 282 - 298. 1946.

(17) Goulden, C. H. Methods of statistical analysis. New York, John Wiley and Sons, Inc. 1939.

(18) Harris, T. E. Note on differentiation under the expectation sign in the fundamental identity of sequential analysis. Ann. Math. Stats. 18: 294. 1947.

(19) Hotelling, H. Note on a matric theorem of A. T. Craig. Ann. Math. Stats. 15: 427 - 429. 1944.

(20) Hsu, P. L. Canonical reduction of the general regression problem. Ann. Eug. Lond. 11: 42 - 46. 1941.

(21) ______ Analysis of variance from the power function standpoint. Biometrika 32: 62 - 69. 1941.

(22) Hsu, P. L. On the power functions of the E^2-test and the T^2-test. Ann. Math. Stats. 16: 278 - 286. 1945.

(23) Irwin, J. O. Mathematical theorems involved in analysis of variance. J. R. S. S. 94: 284 - 300. 1931.

(24) ________ On the independence of the constituent items in the analysis of variance. Supp. J. R. S. S. 1: 236 - 251. 1934.

(25) ________ On the distribution of a weighted estimate of variance and on analysis of variance in certain cases of unequal weighting. J. R. S. S. 105: 115 - 118. 1942.

(26) Johnson, N. L. and Welch, B. L. Applications of the non-central t-distribution. Biometrika 32: 362 - 389. 1940.

(27) Kolodzieczyk, St. On an important class of statistical hypotheses. Biometrika 27: 161 - 190. 1935.

(28) Lehmann, E. H. and Stein, C. Most powerful tests of composite hypotheses. I. Normal Distributions. Ann. Math. Stats. 19: 495 - 516. 1948.

(29) Lehmer, E. Inverse tables of probabilities of errors of the second kind. Ann. Math. Stats. 15: 388 - 398. 1944.

(30) Madow, Wm. G. The distribution of quadratic forms in non-central normal random variables. Ann. Math. Stats. 11: 100 - 103. 1940.

(31) Paulson, E. A note on the efficiency of the Wald sequential test. Ann. Math. Stats. 18: 447 - 449. 1947.

(32) Samuelson, Paul A. The exact distribution of continuous variables in sequential analysis. Econometrica 16: 191 - 198. 1948.

(33) Scheffe, H. On the ratio of the variances of two normal populations. Ann. Math. Stats. 13: 371 - 388. 1942.

(34) Snedecor, G. W. Statistical methods. 4th ed. Ames, Iowa. Collegiate Press, Inc. 1946.

(35) Statistical Research Group, Columbia University. Sequential analysis of statistical data: Applications. New York, Columbia University Press. 1945.

(36) Stein, C. A note on cumulative sums. Ann. Math. Stats. 17: 498 - 499. 1946.

(37) ________ and Wald, A. Sequential confidence intervals for the mean of a normal distribution with known variance. Ann. Math. Stats. 18: 427 - 433. 1947.

(38) Stockman, C. M. and Armitage, P. Some properties of closed sequential schemes. Suppl. J. R. S. S. 8: 104 - 112. 1946.

(39) Tang, P. C. The power function of the analysis of variance tests with tables and illustrations of their use. Stat. Res. Mem., London Univ., 2: 126 - 149. 1938.

(40) Wald, A. On the power function of the analysis of variance test. Ann. Math. Stats. 13: 434 - 439. 1942.

(41) ———— On cumulative sums of random variables. Ann. Math. Stats. 15: 283 - 296. 1944.

(42) ———— Sequential tests of statistical hypotheses. Ann. Math. Stats. 16: 117 - 186. 1945.

(43) ———— Some generalisations of the theory of cumulative sums of random variables. Ann. Math. Stats. 16: 287 - 293. 1945.

(44) ———— Differentiation under the expectation sign in the fundamental identity of sequential analysis. Ann. Math. Stats. 17: 493 - 497. 1946.

(45) ———— Foundations of a general theory of sequential decision functions. Econometrica 15: 279 - 313. 1947.

(46) ———— Sequential analysis. New York, John Wiley & Sons, Inc. 1947.

(47) ———— An essentially complete class of admissible decision functions. Ann. Math. Stats. 18: 549 - 555. 1947.

(48) ———— and Wolfowitz, J. Optimum character of the sequential probability ratio test. Ann. Math. Stats. 19: 326 - 339. 1948.

(49) Wilks, S. S. The large sampling distribution of the likelihood ratio for testing composite hypotheses. Ann. Math. Stats. 9: 60 - 62. 1938.

(50) ——— Mathematical statistics. Princeton, Princeton University Press. 1943.

(51) Wishart, J. A note on the distribution of the correlation ratio. Biometrika 24: 441 - 456. 1932.

(52) Wolfowitz, J. On sequential binomial estimation. Ann. Math. Stats. 17: 489 - 492. 1946.

(53) ——— The efficiency of sequential estimates and Wald's equation for sequential processes. Ann. Math. Stats. 18: 215 - 230. 1947.

VI. ACKNOWLEDGEMENTS

The author wishes to acknowledge the assistance of Dr. G. W. Brown and Dr. A. M. Mood who suggested this problem and acted as advisors in its initial stages. A review of the problem by Dr. G. R. Seth has been valuable in pointing out necessary revisions and additions.